Title

Analysis of resistance-associated substitutions in acute hepatitis C virus infection by deep sequencing across six genotypes and three continents.

Journal

Journal of Viral Hepatitis, 24(1)

Authors

Eltahla, A
Rodrigo, C
Betz-Stablein, B
et al.

Publication Date

2017

DOI

10.1111/jvh.12615

Peer reviewed

HHS Public Access

Author manuscript

J Viral Hepat. Author manuscript; available in PMC 2019 March 17.

Published in final edited form as:

J Viral Hepat. 2017 January ; 24(1): 37–42. doi:10.1111/jvh.12615.

Analysis of resistance-associated substitutions in acute hepatitis C virus infection by deep sequencing across six genotypes and three continents.

Auda A. Eltahla[1], **Chaturaka Rodrigo**[1], **Brigid Betz-Stablein**[1], **Jason Grebely**[2], **Tanya Applegate**[2], **Fabio Luciani**[1], **Janke Schinkel**[3,4], **Gregory J. Dore**[2], **Kimberly Page**[5], **Julie Bruneau**[6], **Meghan D. Morris**[7], **Andrea L. Cox**[8], **Arthur Y. Kim**[9], **Naglaa H. Shoukry**[6], **Georg M. Lauer**[9], **Lisa Maher**[2], **Margaret Hellard**[10,11,12], **Maria Prins**[3,4], **Andrew R. Lloyd**[1], and **Rowena A. Bull**[1] **InC3 Study Group**

[1]School of Medical Sciences, Faculty of Medicine, UNSW Australia, Sydney, NSW, Australia, [2]The Kirby Institute, UNSW Australia, Sydney, NSW, Australia, [3]Department of Internal Medicine, Division of Infectious Diseases, Tropical Medicine and AIDS, Center for Infection and Immunity Amsterdam, Academic Medical Center, Meibergdreef, Amsterdam, The Netherlands, [4]GGD Public Health Service of Amsterdam, Amsterdam, The Netherlands, [5]Department of Internal Medicine, University of New Mexico, Albuquerque, NM, USA, [6]CRCHUM, Université de Montréal, Montreal, QC, Canada, [7]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, USA, [8]Department of Medicine, Johns Hopkins Medical Institutions, Baltimore, MD, USA, [9]Harvard Medical School, Boston, MA, USA, [10]Burnet Institute, Melbourne, VIC, Australia, [11]Alfred Hospital, Department of Infectious Diseases, Melbourne, Australia, [12]Department of Epidemiology and Preventive Medicine, Monash University, Australia.

Abstract

Several direct-acting antivirals (DAA) have now been approved for the treatment of chronic HCV infections, opening the door to imminent interferon-free regiments containing molecules targeting the viral protease (NS3), phosphoprotein (NS5A) and the polymerase (NS5B). Resistance-associated variants (RAVs) have been reported both in treatment naïve patients, and upon treatment using NS3, NS5A and NS5B inhibitors. The prevalence of naturally occurring RAVs in untreated HCV infected subjects has mostly been analysed in subjects infected with GT1 HCV, in the late phase of infection, and only within limited regions of the genome. Therefore, the prevalence of RAVs in non-GT1 infections, and their co-occurrence across the genome remains poorly characterized. In this study, we used next-generation sequencing to analyse full-length HCV genomes for the prevalence of RAVs in people with acute HCV infections from nine international cohorts. RAV were analysed in 180 patients infected with all six major HCV

Correspondence: Dr Rowena A. Bull, School of Medical Sciences, Faculty of Medicine, University of New South Wales, Sydney 2052, Australia, Tel: +61 9385 3838, Tel: +61 9385 1483, r.bull@unsw.edu.au.

Author contribution:

All authors contributed to the design of the InC3 Study. All authors have contributed data to the InC3 Study. Authors AAE and CR completed the laboratory work and the data analysis. The first manuscript was drafted by AE, CR FL, AL and RAB. All authors contributed to and have approved the final manuscript

Conflicts of interest: The authors have no conflicts of interest to disclose.

genotypes (GT1-GT6), and the geographic distribution of RAVs was assessed in 107 GT1a and GT3a samples. While RAVs were detected at varying frequencies across the three genes, and between genotypes, RAVs to multiple DAAs in leading IFN-free regiments were rarely detected in combination. Low-frequency RAVs were also shown to have a GT-specific distribution. The main RAVs with geographic associations were NS3 Q80K in GT1a samples and NS5B M142T in GT3a.

Keywords

hepatitis C virus; inhibitors; direct-acting antivirals; antiviral resistance

Introduction

Hepatitis C virus (HCV) is a significant cause of morbidity and mortality worldwide (1). With a global prevalence of 80–140 million active infections (2, 3), HCV is the leading cause of progressive chronic liver disease, potentially resulting in cirrhosis, liver failure hepatocellular carcinoma and death (3). HCV also exhibits high genetic diversity, both within and between hosts, and has been classified into seven genotypes (GT1-GT7) and 67 confirmed subtypes (4). In addition to this diversity, following transmission of the virus and establishment of infection, the viral variants replicate and generate a genetically diverse population within the host, often referred to as the quasispecies (5).

Historically, treatment options for HCV infections have relied heavily on interferon-α (IFN) and ribavirin (RBV), which are poorly tolerated and variable in terms of success rates (6). In recent years, treatment for HCV has undergone a major transformation with the arrival of direct-acting antivirals (DAAs); molecules that specifically target the viral protease (non-structural (NS) 3, phosphoprotein (NS5A) and the polymerase (NS5B) to inhibit viral replication. Several DAAs have been approved for the treatment of chronic HCV infections, opening the door to widespread uptake of well-tolerated highly effective IFN-free regimens [reviewed in (7)].

In vitro analysis of the effectiveness of these DAAs has revealed that, in most cases, a single amino acid substitution within the target viral protein can cause a significant loss of antiviral potency. Likewise, such resistance-associated variants (RAVs) have been reported upon single agent treatment of patients infected with HCV using NS3, NS5A and NS5B inhibitors, and have been associated with treatment failure (8). While in most cases, such variants carry a significant fitness cost to the virus and disappear after treatment cessation, they can persist for up to two years after treatment (9, 10), therefore potentially affecting re-treatment options (11), and also being transmitted to new hosts (12). In addition, the high diversity of HCV has meant many viral isolates naturally carry residues which confer resistance to some approved DAAs as well as those in development, this natural resistance is highly genotype/subtype specific (13, 14). Such naturally occurring RAVs have been shown to affect treatment response rates both in IFN-based and IFN-free regimens (reviewed in (15)).

Several studies have analysed the prevalence of HCV RAVs in treatment-naïve populations. These studies however were limited by sample size, variant detection sensitivity and poor

representation of non-GT1 HCV infections. Furthermore, studies have so far focused on samples collected in the course of chronic infection and from limited geographic regions. Therefore, the potential transmission of RAVs and their natural prevalence in non-GT1 infections remains poorly characterized. In addition, the co-occurrence of low frequency variants across NS3, NS5A and NS5B within a single infection remains unclear. This is particularly important as the landscape of HCV therapies is moving towards combination DAAs targeting different viral proteins simultaneously.

The aim of this study was to perform a comprehensive, ultra-deep analysis of the prevalence of RAVs in samples from patients acutely infected with HCV. We compare the prevalence of these RAVs across all six major genotypes, and the relative distribution of RAVs across three continents.

Materials and Methods:

Study population and design

Participants included in this study were selected from nine prospective cohorts of high-risk individuals recruited and followed between 1985 and 2014, which formed the International Collaboration of Incident HIV and Hepatitis C in Injecting Cohorts (InC3), described previously (16). Samples were selected from the repositories of the individual InC3 cohorts from North America (n=4), Australia (n=4) and Europe (n=1) (Table 1). To be included in this study, only acute phase samples were selected (i.e. collected within 180 days of the estimated date of infection, calculated as described previously) (17). Only samples with a viral load greater than 1,000 IU/ml were included, and all participants were treatment naïve at the time of collection. Ethics approval for this study was granted by the University of New South Wales, Australia (No. 14201). In addition, each of the individual cohorts had ethical approval for data and specimen collections. All data pertaining to individual participants were de-identified.

RNA extraction, amplicon generation and sequencing

Viral RNA was extracted from plasma samples and amplicons covering the HCV genome were generated either as single near-full length products, or as three hemigenomes, as described previously (18). Due to primer-design limitations, samples with GT4a, GT4d, and GT5a generated amplicons that only extended to genome position 9,277 (with reference to GT1a strain H77), and therefore NS5B RAVs past this position were not analysed. Amplicons were sequenced either using the Roche 454 FLX (n=10), or an Illumina MiSeq Benchtop (n=172) sequencing platform. Roche 454 FLX sequencing was performed as described previously (18). Illumina sequencing libraries were prepared using either the Nextera XT or TruSeq Nano DNA Library Prep Kits (Illumina) according to the manufacturer's instructions.

Sequence alignment and SNP detection

Sequence read quality was checked using FastQC (Babraham Institute, http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and trimmed using Geneious package version 8 (19), setting the error rate to 0.3% (equivalent to QC25). Reads were then aligned

using Bowtie 2 (20) against the corresponding genotype reference sequence (GenBank accession numbers shown in brackets): GT1a (NC_004102), GT1b (AJ238799), GT2a (AB047639), GT2b (AB030907), GT3a (HPCEGS), GT4a (Y11604), GT4d (DQ418786), GT4m (FJ462433), GT5a (AF064490), GT6a (Y12083), GT6I (EF424628). For the purpose of detecting nucleotide variants resulting in DAA RAVs, reference sequences were modified to carry DAA-sensitive residues at the sites analysed in this study. Single-nucleotide polymorphism (SNP) analysis was performed using the Geneious package with minimum variant frequency threshold of 0.001, maximum variant P-value of 10^{-6}, and a minimum coverage of 1,000. For 454 sequences, a coverage threshold of 400 was chosen to account for the lower coverage relative to Illumina sequences.

Phylogenetic analysis

Consensus sequences generated from all samples included in the study were aligned using MUSCLE and a phylogenetic tree was constructed using the Neighbour-joining method implemented in Geneious package version 8 (19).

Statistical analysis

To investigate whether specific RAVs are more prevalent in certain regions or genotypes, Fisher's exact test of independence was employed. Results were considered statistically significant if the corresponding p value was < 0.05. Analysis by GT included GT1a (n=74) and GT3a (n=66). Analysis by region was performed to compare RAVs prevalence in Australia and North America stratified for GT1a and GT3a (Table 1). Whether a participant carried a RAV within their viral population was defined as a variant frequency greater than the cut-off of either 10% or 1%.

Results

Study population

In total, NGS data was generated for 180 samples collected from cohort participants during acute HCV infection between 2000 and 2014. The mean age of these participants was 26.5 (standard deviation 8.1 years), 69% were male, 5.6% were co-infected with HIV. All participants were naïve to DAA-based therapy. The majority of participants were infected with HCV GT1a (41.8%, n=76) and GT3a (36.3%, n=66). A minority of participants were also infected with HCV GT1b (n=5), GT2a (n=2), GT2b (n=12), GT4a (n=7), GT4d (n=5), 4m (n=1), GT5a (n=4), GT6a (n=3) and 6I (n=1). The distribution of HCV genotypes stratified by country of origin is shown in Table 1.

Resistance variants within the NS3 coding region

The prevalence of drug-resistant variants was calculated from the deep sequence data that was generated across NS3, NS5A and NS5B for each participant. A summary of the frequency for the analysed RAVs in the entire dataset is summarised in Figure 1. It is worth noting that the prevalence of analysed RAVs in the study population varies depending on the frequency threshold set within the viral quasispecies for each study. In order to allow comparisons between our dataset and other studies, the prevalence of each RAV was assessed with a range of frequency thresholds (Figure S1). Unsurprisingly, and similar to

previous reports, the frequency of RAVs was less variable when the threshold was set to 10% (Figures S1–S3), hence this was taken as the frequency threshold for reporting RAV prevalence.

In regards to protease inhibitors (PI) specifically, the most prevalent RAV was L175. This variant, known to be associated with the first-generation PI boceprevir, was the dominant variant in all genotypes with the exception of GT1b and GT6a (Figure 1). On the other hand, variations at NS3 position 36 were only observed at low frequencies (<2%) in GT1 and GT6 sequences, and the consensus-level variant L36 was observed in GT2-GT5 sequences (Table 2). Q80K is a known NS3 variant associated with resistance to multiple protease inhibitors including simeprevir, paritaprevir, and asunaprevir. Q80K was detected in the dataset at frequencies >10% in 22/74 GT1a samples (30%, Table 2). Q80K was also found as the consensus residue for all GT5a and GT6a, but not for GT6I (noting that only one GT6I sample was analysed here). Variants at this position were rarely observed in GT2b samples, even at low frequency (i.e. less than 1%) (Figure 1).

The variant S122R, known to be associated with reduced activity of simeprevir (21) was only detected in GT2b subtypes. Interestingly, low-frequency variation S122R variants were rarely detected in non-GT2b sequences, including the two GT2a samples (Figure 1). First-generation PI RAVs T54A/S, and those conferring resistance to multiple PIs in positions V55 and R155 were only detected in a minority of GT1a sequences at frequencies >10% (Table 2 and Figure S1), whereas F43 and V132 RAVs were only detected in 1 of 33 GT3a samples each (Figure 1 and Table 2). V132 was also detected at consensus-level in 2 of 5 GT1b sequences. NS3-Y56 RAVs, which affect DAAs paritaprevir and asunaprevir, were also rarely detected across all genotypes, but were more common at low-frequency in GT1a samples, and absent from all GT5a samples.

Variations were also rarely detected in Q41, S138, V158 and V170. Variations at 107I, associated with resistance to telaprevir were only detected at consensus-level in the one GT6I sample, and below 10% frequencies in all non-GT1 genotypes (Figure 1 and Figure S1). Intriguingly, nearly all (84%) samples included in the study, with the exception of GT4a samples, carried the low-frequency variants 156T/V, which are associated with resistance to a number of PIs. NS3 RAV Q168 was the dominant residue in all but one GT3a sequence, which instead carried H168 (Table 2).

In order to analyse the geographic distribution of NS3 RAVs, samples obtained from North America and Australia were compared for GT1a and GT3a. The only NS3 RAV with a significant geographic difference was Q80K (Table 5). The prevalence of Q80K was lower among participants with GT1a from Australia (1/25, 4%) compared to participants with GT1a from North America (20/48, 42%, $P<0.001$). Additionally, one North American sequence carried the Q80K variant at lower frequencies (25%, Figure 1). Interestingly, this participant also carried another RAV at the same position, R80, at low frequencies (12%), and additional low-frequency RAVs at several other positions, both within and outside the NS3-coding region (Figure 1).

Resistance variants within the NS5A coding region

NS5A RAVs at position 24 and 26 were rarely detected at frequencies >10% in all genotypes, even at low frequency (Figure S2). Only one GT1a participant carried R24 at 44% and one GT3a participant had E26 at 26% frequency (Figure 1). Whereas lower frequency variants were detected in other genotypes, none could be detected in GT1b and GT2a/2b sequences.

Resistance to a number of NS5A inhibitors, including ombitasvir, daclatasvir and ledipasvir has been associated with variations at NS5A position 28, which were detected across genotypes. The variant V28, implicated in resistance to NS5A inhibitor, ombitasvir was detected at the consensus level in 2 of 74 (3%) GT1a samples, and in the GT6I infected sample. The variant was also present at lower frequencies in another GT1a sample and a GT3a sample (18–49% frequencies, Figure 1 and Table 2). Other variants such as T28 which confer resitance to all approved NS5A inhibitors were rarely detected at frequencies >10% (Table 3 and Figure S2). Low-frequency RAVs at this position could not be detected in any of the GT1b and GT5a samples, but were present at different levels in all other GTs, and the most relevant variants, 28T/V were limited to GT1a and GT3a samples.

Similar to variants within NS5A residue 28, substitutions at position 30 were also associated with a number of NS5A DAAs, and are characterised by consensus-level variation between genotypes (Table 3). In GT1a participants, 2 of the 48 (4%) samples derived from North America, and none of the Australian-derived sequences, carried the variant H30 (Table 3). Similarly, low prevalence of RAVs 30K/T were detected for GT3a (Figure 1). No low-frequency RAVs at position 31 were detected in GT1a/1b, GT5 and GT6 sequences. Only RAV 31F could be detected at low frequencies in GT3a samples, while GT2 and GT4a samples, in addition to consensus-level 31M variants, carried low-frequency (<2%) 31V/I. Similarly, while no RAVs were detected at the consensus level at NS5A 32, low frequency variants, <2%, were detected predominantly in GT1 and GT2 samples (Figure 1). NS5A 32 and 38 substitutions were only detected at very low frequencies (<2%) across genotypes.

Variations at NS5A position 93 which confer resistance to a number of RAVs were also rarely detected in our dataset, with only one GT1a sequence carrying 93S, and only one GT1a and one GT3a sample with 93H at the consensus-level, while another GT3a and a GT4a sample carried this RAV at medium-level frequencies (39% and 57%, respectively). Low frequency RAVs 92T were predominantly detected in GT1 and GT4-GT6. No significant geographic differences were observed for NS5A RAVs in our dataset with either GT1a or GT3a (Table 5).

Resistance variants within the NS5B coding region

In agreement with previous reports, RAVs associated with the NNI sofosbuvir (positions 282, 289, 159, 320 and 321) were rarely detected in our dataset. S282T was not detected in any of the sequences, while L159F was only detected in one GT1a sample, which carried this RAV as the dominant variant. S15G was detected only in GT2b samples. This RAV has been associated with resistance to some HCV NIs, but only in combination with C223H/V321I (22), and this co-occurence was not present in any sample; RAVs at 223 were not

detected, while only low-frequency variants were detected at 321, and were primarily alanine residues. Substitutions at M289, which have been implicated in resistance to NIs such as sofosbuvir, were detected as majority variants in two GT2b and the one GT6I sample (Table 4), and as minority variants in all non-GT1 participants (Figure 1). It is worth noting that M289L was primarily observed in GT2a, but not GT1a/1b replicons, in the preclinical assessment of sofosbuvir, and while this mutation had a minor effect on drug potency, in combination with S282T there was an additive effect (23). In our dataset, however, only 289I was detected in GT2b (Table 4). The NS5B RAV C316Y which has been reported to confer resistance to different classes of NIs and NNIs, was rarely observed at frequencies >2%, whereas minority variants were detected across all GTs (Figure 1).

Resistance to the recently approved NNI dasabuvir has been associated with variants at C316, S368, M411, M414, Y448, A553, G554 and S556 positions. In the GT1a participants, only C316Y (1/74), A553V (2/74), G554D (1/74) and S556G (4/74) were detected in frequencies >10%, and only S556G variants were detected at the consensus-level (Figure 1 and Figure S3). Of the GT1b samples, 1 of 5 (20%) carried the S556G variant at a frequency of 42%. With the exception of position 368, lower-frequency variants associated with dasabuvir were present in all GT1 samples. Only GT6 samples did not contain consensus-level RAVs at these positions. GT4 samples carried 414 whereas other GTs carried subtype-specific variants at 414L/V/I. RAVs V553 and G556 were also present at the consensus-level in different subtypes (Table 4). Of the NS5B substitutions associated with resistance to beclabuvir, A421V was frequently observed at a frequency >10% both in Australian (25%) and North American (10%) derived GT1a samples, and almost all GT1a samples carried this RAV at lower frequencies (0.2–4%). On the other hand, beclabuvir-associated RAVs at position 495 were only detected at frequencies up to 2.2% (Figure 1). Consensus-level V421 was detected in all non-GT1 sequences (Table 4). Variants associated with a class of NNIs binding to the thumb II site, including 419I, 423I, 482L and 494A were genotype and subtype-specific (Figure 1) and with the exception of GT1b samples, >97% of all non-GT1b samples had consensus level NS5B V499A (Table 4).

When the geographic distribution of NS5B RAVs was analysed, N142T, an NS5B mutation that has been associated with resistance to some HCV NIs but not sofosbuvir, was only detected in GT3a samples derived from North America; 4 of 33 (15%) participants carried this variant as a dominant variant, while no minor variants were detected in any of the other samples (Figure 1). While this RAV was significantly associated with GT3a ($P = 0.02$), its geographic association with North American sampling fell below the significance level ($P = 0.053$, Table 5).

Resistance to advanced IFN-free combinations

Analysis of the entire coding region for NS3, NS5A and NS5B in this study allowed for the simultaneous analysis of RAVs involved in resistance to upcoming regimens containing molecules targetting multiple HCV proteins. As mentioned, no sofosbuvir-associated RAVs were detected in our GT1 dataset, and therefore combinations of RAVs affecting double-therapy regimens containing sofosbuvir and either ledipasvir, daclatasvir, and velpatasvir were not detected.

RAVs relevant to other double and triple therapy regimens which have advanced through clinical trials have been analysed for GT1a sequences and are highlighted in Figure 1. These regimens include i) ombitasvir, paritaprevir and dasabuvir; ii) daclatasvir, asunaprevir, and beclabuvir; and iii) grazoprevir and elbasvir. Of the RAVs relevant to the recently licensed regimen containing ombitasvir, paritaprevir and dasabuvir (24, 25), no samples contained resistant variants to all three molecules in the regimen, however NS3 K80 was detected with NS5B V553, with NS5B D554, with NS5B V553/G556, each in one sample. NS5A/NS5B variants detected in combination were NS5A V28 with NS5B G556, and were detected in 2 of 50 (4%) samples. In relation to another IFN-free triple regimen containing daclatasvir, asunaprevir and beclabuvir (26), six participants carried the NS3 Q80K in combination with NS5B 421V (Figure 1). Of these participants, three also carried the NS5A H30 variant. Interestingly, one sequence carried NS3 Q80K, Q80R, NS5A H30, and NS5B V421 all at frequencies 12–28% of the viral population (Figure 1). At frequencies >10%, the NS3 RAV Q80K was detected in GT1a sequences in combination with NS5A M28V/T in two samples, with NS5A R30H in three samples, and with NS5A Y93S in yet another sample (Figure 1). While low-frequency variants were detected in combination across NS3 156 and NS5A 28 sites, which are relevant to multiple IFN-free regimens (26, 27), these were not detected in combination at frequencies >10% of the viral population. Similarly, variants associated with resistance to the combination regiment grazoprevir/elbasvir (27) were not detected in combination in any GT1 samples.

Discussion:

The rapid evolution of HCV therapeutics has meant that there is a need for a better understanding of the mutations encoding for resistance against antiviral agents in the clinical setting. The aim of this study was to understand the prevalence of such RAVs during the acute phase of HCV infection, and how these RAVs differ across genotypes and between geographic regions.

While RAVs could have variable impact on the drug potency in different HCV genotypes (15), the large number of samples representing the major GTs in this study has allowed the direct comparison of this prevalence particularly between GT1a and GT3a (Tables 2–4). Besides consensus-level variations, notable differences between genotypes were found to be NS3 Q80K, associated with GT1a samples, and NS5B M142T, associated with GT3a. The difference in RAV prevalence between HCV strains has so far been largely confined to GT1a/GT1b, and under-reported for other genotypes (28). Understanding these genotypic differences is crucial for future therapies where pan-genotypic regimens are the ultimate goal, and where the necessity for monitoring resistance before or after treatment remains to be resolved.

Our analysis has also demonstrated a significant difference in the prevalence of some RAVs within the same genotype, but across different geographic regions (Table 5). The geographic spread of NS3 Q80K has been well documented, reaching levels of 48% in North American GT1a isolates, 19% in Europe and 9% in South America (29). Furthermore, in a recent study of the European region, variable rates were reported for Q80K prevalence across different countries, ranging between 0% in Bulgaria to around 18% in the UK (30). In agreement with

previous reports, the prevalence of this RAV in the North American continent samples studied here matched those previously reported for North America in GT1a (38%). However, only 4% of GT1a sequences of Australian origin carried this variant. This RAV has been associated with resistance to a number of approved NS3 inhibitors including simeprevir and paritaprevir. Interestingly, tracing the evolutionary history of this polymorphism has revealed that the majority of isolates carrying this variant originated from a single lineage that was estimated to have arisen during the 1940s (31). Combined with our data, this indicates that DAA regimens affected by baseline Q80K variants are more likely to be successful in the Australian population.

Within-host, the frequency at which baseline RAVs become clinically relevant remains unresolved, and is likely to be dependent on the treatment regimen used (15, 32). For HIV-infected patients, minority variants not detected by population sequencing could significantly affect the patients' response to antiretroviral therapies (33). Our analysis of this large dataset using deep-sequencing platforms revealed not only the consensus-level variation across known RAV sites, but also the prevalence of low-level (<10%) RAVs across the major genotypes. For instance, in agreement with previous reports, resistance variants S282T and L159F, associated with resistance to HCV NIs such as sofosbuvir were rarely observed in our dataset, even at low frequency, which is explained by the impaired replicative fitness of this variant (34). In contrast, NS3 A156T/V variants, which are relevant to a number of PIs, were carried at low frequencies in the majority of samples, with the exception of those belonging to GT4a HCV. This is despite the apparent fitness cost these variants confer to GT1a/1b *in vitro* (35). Similarly, NS5A M28T/V was detected at low frequencies in a large number of samples (Figure 1). It should be noted, however, that the absence of these variants in GT1b samples could be explained by the loss of replication efficiency of GT1b, but not GT1a, viruses carrying NS5A 28 RAVs in cell culture (36). Other variants which had GT-specific prevalence were NS3 Q80, which could not be detected in GT2b, while S122R was rarely detected in non-GT2b samples. Similarly, a notable absence of the following RAVs was observed in specific genotypes; NS3 Y56H/I in GT5a, NS3 R155 RAVs in GT4a, NS5A L31I/V in GT1a/1b, 5a and 6a and NS5B M414T/I/L/V in GT2 samples (Figure 1). It is known that the genetic barrier to resistance not only differs between classes of DAAs (13, 28), but also between different HCV subtypes towards the same DAA (13). This is believed to be attributable to the number and type of nucleotide changes required to generate these RAVs, and the fitness cost of these RAVs in different genotypes (37). These genotypic differences in tolerance to even low frequency variations that are associated with DAA resistance may provide a valuable insight into which molecules could be targeted for pan-genotypic regimens, and which genotypes might need screening for antiviral resistance before or while on therapy.

Conclusion:

This is the largest study assessing the prevalence of HCV-related resistance variants in acute infection to characterise whether these RAV are commonly associated with transmitted variants. It also provides the most comprehensive estimate of the prevalence of HCV RAVs across the HCV genome for six different genotypes across three continents using deep

sequencing of the HCV genome, which is needed if these emerging therapies are to be applied on a global scale.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

The authors acknowledge the support of the Ramaciotti Centre for Genomics, UNSW, Australia in conducting the next generation sequencing. We are grateful to the study participants and to field and laboratory staff (Suzy Teutsch, Luke McCredie, Geng Zang, Rachel Bouchard, Elisabeth Deschenes, Marie-Eve Turcotte. Yanie Nan, Mélanie Gagnon, Isabelle Boisvert, Nathalie Bédard, Rachel Sack Davis, Paul Dietze, Peter Higgs, Campbell Aitken) and key co-investigators (Elise Roy and Didier Jutras-Aswad).

Financial support: The infrastructure for sharing of data and specimens in InC^3 was funded by the National Institute on Drug Abuse (NIDA) R01DA031056. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of NIDA or the National Institutes of Health (NIH). Research support for the InC^3 cohorts includes: the Netherlands National Institute for Public Health and the Environment to the Amsterdam Cohort Study (ACS); Baltimore Before and After Study (BBA) - National Institutes of Health (NIH U19 AI088791); Boston Acute HCV Study: transmission, immunity, outcomes network (BAHSTION) is funded under NIH NIAID U19 U19 AI066345; ATAHC – NIDA RO1 DA 15999–01; HITS-p - National Health and Medical Research Council of Australia (NHMRC) - Project No. 222887, Partnership No. 1016351, Program Nos. 510488 and 1053206; HITS-c – UNSW Hepatitis C Vaccine Initiative and NHMRC Project Grant No. 630483; Networks/MIX – NHMRC Project Grants Nos. 331312 and 545891 and the Victorian Operational Infrastructure Support Programme (Department of Health, Victoria, Australia); HepCo - the Canadian Institutes of Health Research (MOP-103138 and MOP-106468; Réseau SIDA Maladies Infectieuses, Fonds de la Recherche du Québec-Santé); UFO – NIH R01 DA016017. The following InC^3 investigators are funded by research fellowships: NHMRC, Australia – JG, MH, LM, GD, ARL; Fonds de la Recherche du Québec, Canada – JB and NHS.

Abbreviations:

HCV	Hepatitis C virus
DAA	direct-acting antivirals
RAV	resistance-associated variants
IFN	interferon
RBV	ribavirin
InC3	International Collaboration of Incident HIV and Hepatitis C in Injecting Cohorts
HIV	Human immunodeficiency virus
BWA	Burrows-Wheeler Alignment tool
MUSCLE	MUltiple Sequence Comparison by Log-Expectation
RNA	Ribonucleic acid
IU	International Units
PCR	Polymerase chain reaction

NGS	Next generation sequencing
NS	non-structural

References:

1. Hajarizadeh B, Grebely J, Dore GJ. Epidemiology and natural history of HCV infection. Nature Reviews Gastroenterology and Hepatology 2013;10:553–562. [PubMed: 23817321]
2. Gower E, Estes C, Blach S, Razavi-Shearer K, Razavi H. Global epidemiology and genotype distribution of the hepatitis C virus infection. Journal of hepatology 2014;61:S45–S57. [PubMed: 25086286]
3. Mohd Hanafiah K, Groeger J, Flaxman AD, Wiersma ST. Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence. Hepatology 2013;57:1333–1342. [PubMed: 23172780]
4. Smith DB, Bukh J, Kuiken C, Muerhoff AS, Rice CM, Stapleton JT, Simmonds P. Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment web resource. Hepatology 2014;59:318–327. [PubMed: 24115039]
5. Martell M, Esteban J, Quer J, Genesca J, Weiner A, Esteban R, Guardia J, et al. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. Journal of Virology 1992;66:3225–3229. [PubMed: 1313927]
6. Ghany MG, Strader DB, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C: an update. Hepatology 2009;49:1335–1374. [PubMed: 19330875]
7. Asselah T, Boyer N, Saadoun D, Martinot-Peignoux M, Marcellin P. Direct-acting antivirals for the treatment of hepatitis C virus infection: optimizing current IFN-free treatment and future perspectives. Liver International 2016;36:47–57. [PubMed: 26725897]
8. Poveda E, Wyles DL, Mena Á, Pedreira JD, Castro-Iglesias Á, Cachay E. Update on hepatitis C virus resistance to direct-acting antiviral agents. Antiviral research 2014;108:181–191. [PubMed: 24911972]
9. McPhee F, Hernandez D, Yu F, Ueland J, Monikowski A, Carifa A, Falk P, et al. Resistance analysis of hepatitis C virus genotype 1 prior treatment null responders receiving daclatasvir and asunaprevir. Hepatology 2013;58:902–911. [PubMed: 23504694]
10. Dvory-Sobol H, Wyles D, Ouyang W, Chodavarapu K, McNally J, Cheng W, Shafran S, et al. Long-term persistence of HCV NS5A variants after treatment with NS5A inhibitor ledipasvir. J Hepatol 2015;62:S221.
11. Krishnan P, Tripathi R, Schnell G, Reisch T, Beyer J, Dekhtyar T, Irvin M, et al. O057: Long-term follow-up of treatment-emergent resistance-associated variants in NS3, NS5A and NS5B with paritaprevir/r-, ombitasvir-and dasabuvir-based regimens. Journal of Hepatology 2015;62:S220.
12. Franco S, Tural C, Nevot M, Moltó J, Rockstroh JK, Clotet B, Martinez MA. Detection of a Sexually Transmitted Hepatitis C Virus Protease Inhibitor-Resistance Variant in a Human Immunodeficiency Virus–Infected Homosexual Man. Gastroenterology 2014;147:599–601. e591. [PubMed: 24859204]
13. Eltahla AA, Tay E, Douglas MW, White PA. Cross-genotypic examination of hepatitis C virus polymerase inhibitors reveals a novel mechanism of action for thumb binders. Antimicrob Agents Chemother 2014;58:7215–7224. [PubMed: 25246395]
14. Scheel TK, Gottwein JM, Mikkelsen LS, Jensen TB, Bukh J. Recombinant HCV variants with NS5A from genotypes 1–7 have different sensitivities to an NS5A inhibitor but not interferon-α. Gastroenterology 2011;140:1032–1042. e1036. [PubMed: 21111742]
15. Sarrazin C The importance of resistance to direct antiviral drugs in HCV infection in clinical practice. Journal of hepatology 2015.
16. Grebely J, Morris MD, Rice TM, Bruneau J, Cox AL, Kim AY, McGovern BH, et al. Cohort profile: the international collaboration of incident HIV and hepatitis C in injecting cohorts (InC3) study. Int J Epidemiol 2013;42:1649–1659. [PubMed: 23203695]

17. Hajarizadeh B, Grady B, Page K, Kim AY, McGovern BH, Cox AL, Rice TM, et al. Patterns of hepatitis C virus RNA levels during acute infection: the InC3 study. PLoS One 2015;10:e0122232. [PubMed: 25837807]
18. Bull RA, Luciani F, McElroy K, Gaudieri S, Pham ST, Chopra A, Cameron B, et al. Sequential bottlenecks drive viral evolution in early acute hepatitis C virus infection. PLoS Pathog 2011;7:e1002243. [PubMed: 21912520]
19. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 2012;28:1647–1649. [PubMed: 22543367]
20. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature methods 2012;9:357–359. [PubMed: 22388286]
21. Lenz O, Verbinnen T, Lin T-I, Vijgen L, Cummings MD, Lindberg J, Berke JM, et al. In vitro resistance profile of the hepatitis C virus NS3/4A protease inhibitor TMC435. Antimicrobial agents and chemotherapy 2010;54:1878–1887. [PubMed: 20176898]
22. Lam AM, Espiritu C, Bansal S, Steuer HMM, Zennou V, Otto MJ, Furman PA. Hepatitis C virus nucleotide inhibitors PSI-352938 and PSI-353661 exhibit a novel mechanism of resistance requiring multiple mutations within replicon RNA. Journal of virology 2011;85:12334–12342. [PubMed: 21957306]
23. Lam AM, Espiritu C, Bansal S, Steuer HMM, Niu C, Zennou V, Keilman M, et al. Genotype and subtype profiling of PSI-7977 as a nucleotide inhibitor of hepatitis C virus. Antimicrobial agents and chemotherapy 2012:AAC. 00054–00012.
24. Feld JJ, Kowdley KV, Coakley E, Sigal S, Nelson DR, Crawford D, Weiland O, et al. Treatment of HCV with ABT-450/r–ombitasvir and dasabuvir with ribavirin. New England Journal of Medicine 2014;370:1594–1603. [PubMed: 24720703]
25. Ferenci P, Bernstein D, Lalezari J, Cohen D, Luo Y, Cooper C, Tam E, et al. ABT-450/r–Ombitasvir and Dasabuvir with or without Ribavirin for HCV. New England Journal of Medicine 2014;370:1983–1992. [PubMed: 24795200]
26. Poordad F, Sievert W, Mollison L, Bennett M, Tse E, Bräu N, Levin J, et al. Fixed-dose combination therapy with daclatasvir, asunaprevir, and beclabuvir for noncirrhotic patients with HCV genotype 1 infection. JAMA 2015;313:1728–1735. [PubMed: 25942723]
27. Roth D, Nelson DR, Bruchfeld A, Liapakis A, Silva M, Monsour H, Martin P, et al. Grazoprevir plus elbasvir in treatment-naive and treatment-experienced patients with hepatitis C virus genotype 1 infection and stage 4–5 chronic kidney disease (the C-SURFER study): a combination phase 3 study. The Lancet 2015;386:1537–1545.
28. Eltahla AA, Luciani F, White PA, Lloyd AR, Bull RA. Inhibitors of the Hepatitis C Virus Polymerase; Mode of Action and Resistance. Viruses 2015;7:5206–5224. [PubMed: 26426038]
29. Lenz O, Verbinnen T, Fevery B, Tambuyzer L, Vijgen L, Peeters M, Buelens A, et al. Virology analyses of HCV isolates from genotype 1-infected patients treated with simeprevir plus peginterferon/ribavirin in Phase IIb/III studies. Journal of hepatology 2015;62:1008–1014. [PubMed: 25445400]
30. Sarrazin C, Lathouwers E, Peeters M, Daems B, Buelens A, Witek J, Wyckmans Y, et al. Prevalence of the hepatitis C virus NS3 polymorphism Q80K in genotype 1 patients in the European region. Antiviral research 2015;116:10–16. [PubMed: 25614456]
31. McCloskey RM, Liang RH, Joy JB, Krajden M, Montaner JS, Harrigan PR, Poon AF. Global origin and transmission of hepatitis C virus NS3 Q80 K polymorphism. Journal of Infectious Diseases 2014:jiu613.
32. Wilson EM, Kattakuzhy S, Sidharthan S, Sims Z, Tang L, McLaughlin M, Price A, et al. Successful Retreatment of Chronic HCV Genotype-1 Infection With Ledipasvir and Sofosbuvir After Initial Short Course Therapy With Direct-Acting Antiviral Regimens. Clinical Infectious Diseases 2015:civ874.
33. Van Laethem K, De Munter P, Schrooten Y, Verbesselt R, Van Ranst M, Van Wijngaerden E, Vandamme A-M. No response to first-line tenofovir+ lamivudine+ efavirenz despite optimization according to baseline resistance testing: impact of resistant minority variants on efficacy of low genetic barrier drugs. Journal of clinical virology 2007;39:43–47. [PubMed: 17369083]

34. Lam AM, Espiritu C, Bansal S, Steuer HMM, Niu C, Zennou V, Keilman M, et al. Genotype and subtype profiling of PSI-7977 as a nucleotide inhibitor of hepatitis C virus. Antimicrobial Agents and Chemotherapy 2012;56:3359–3368. [PubMed: 22430955]

35. Pilot-Matias T, Tripathi R, Cohen D, Gaultier I, Dekhtyar T, Lu L, Reisch T, et al. In vitro and in vivo antiviral activity and resistance profile of the hepatitis C virus NS3/4A protease inhibitor ABT-450. Antimicrobial agents and chemotherapy 2015;59:988–997. [PubMed: 25451053]

36. Krishnan P, Beyer J, Mistry N, Koev G, Reisch T, DeGoey D, Kati W, et al. In vitro and in vivo antiviral activity and resistance profile of ombitasvir, an inhibitor of hepatitis C virus NS5A. Antimicrobial agents and chemotherapy 2015;59:979–987. [PubMed: 25451055]

37. Margeridon-Thermet S, Le Pogam S, Li L, Liu TF, Shulman N, Shafer RW, Najera I. Similar prevalence of low-abundance drug-resistant variants in treatment-naive patients with genotype 1a and 1b hepatitis C virus infections as determined by ultradeep pyrosequencing. 2014.

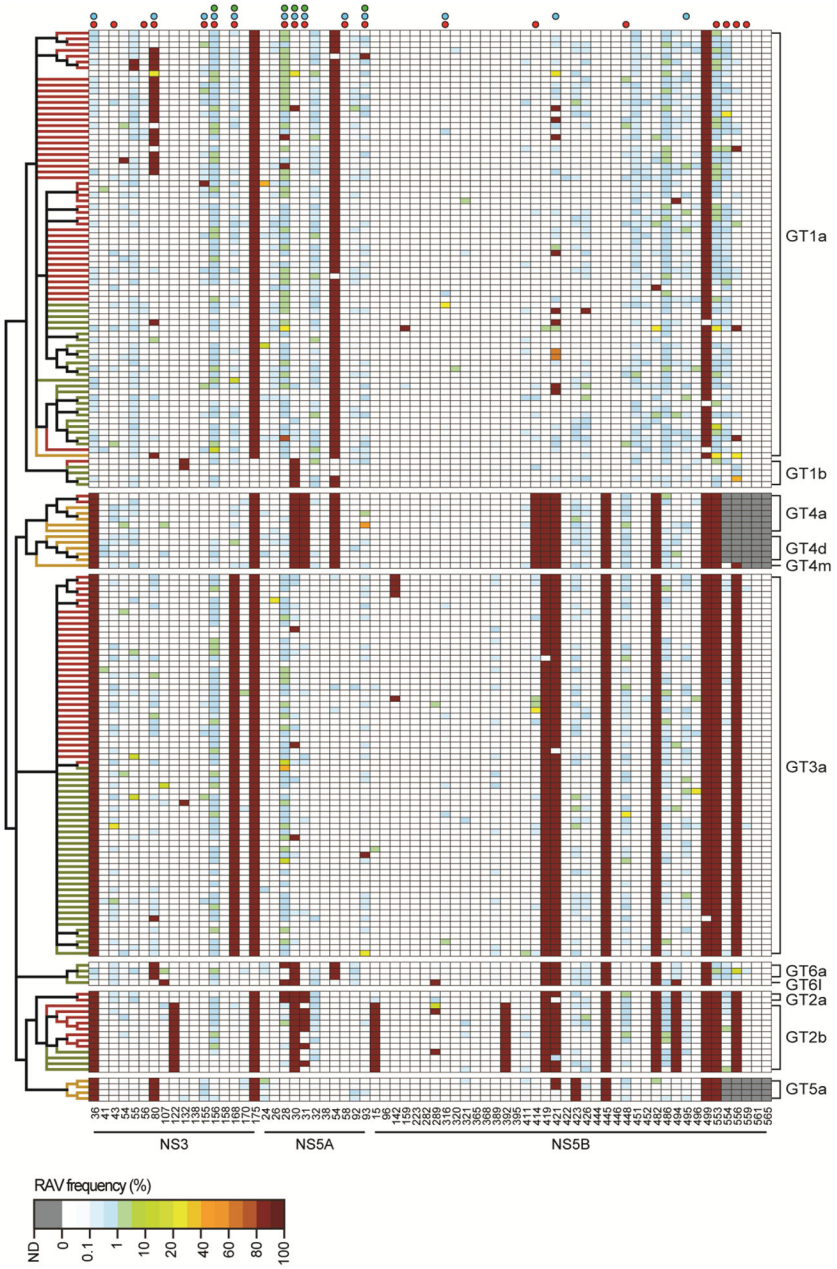

Figure 1. A summary of the resistance-associated variants (RAVs) detected by deep-sequencing across the HCV genome.

RAVs detected by deep-sequencing of the HCV genome are shown together with their frequency of occurrence as a heat map. Rows represent individual samples and columns indicate a position within NS3, NS5A and NS5B which is known to confer resistance to DAAs. A neighbour-joining phylogenetic tree based on the NS5B consensus sequences is also shown. Tree branches are coloured according to the sample origin, where North American samples are shown in red, Australian samples in green and European samples in yellow. The genotype of the sample is indicated by the bracket on the right. The frequency of RAVs is indicated by the color according to the scale bar, and positions where RAVs were not analysed are shown in grey. RAVs associated with new IFN-free DAA combination

regimens are highlighted at the top of the figure; ombitasvir, paritaprevir and dasabuvir (red spheres), daclatasvir, asunaprevir, and beclabuvir (blue spheres), grazoprevir and elbasvir (green spheres).

Table 1

Summary of samples included in the study by genotype and country of origin.

Genotype	Number of Samples (n=180)	Country		
		Australia (n=70)	US/Canada (n=94)	Netherlands (n=16)
GT1a	74 [41.1%]	25	48	1
GT1b	5 [2.8%]	4	1	
GT2a	2 [1.1%]		2	
GT2b	12 [6.7%]	4	8	
GT3a	66 [36.7%]	33	33	
GT4a	7 [3.9%]		2	5
GT4d	5 [2.8%]			5
GT4m	1 [0.6%]			1
GT5a	4 [2.2%]			4
GT6a	3 [1.7%]	3		
GT6I	1 [0.6%]	1		

Table 2.

Prevalence of NS3 resistance associated variants detected in this study*.

		Prevalence of RAVs [%]							
NS3 Position	**Variants Detected**	**GT1a (n=74)**	**GT3a (n=66)**	**GT1b (n=5)**	**GT2b (n=12)**	**GT4a (n=7)**	**GT4d (n=5)**	**GT5a (n=4)**	**GT6a (n=3)**
36	L	-	100	-	100	100	100	100	-
43	S	-	2	-	-	-	-	-	-
54	A	1	-	-	-	-	-	-	-
55	A	3	-	-	-	-	-	-	-
80	K/R	30 (K), 1 (R)	2 (K)	-	-	-	-	100 (K)	100 (K)
122	R	-	-	-	100	-	-	-	-
132	V	-	2	40	-	-	-	-	-
155	K	1	-	-	-	-	-	-	-
168	Q/H	-	2 (H), 98 (Q)	-	-	-	-	-	-
175	L	100	100	-	100	100	100	100	-

* Only variants detected at frequency >10% of the quasispecies are shown

Table 3.

Prevalence of NS5A resistance associated variants detected in this study*.

		Prevalence of RAVs [%]							
NS5A Position	**Variants Detected**	**GT1a (n=74)**	**GT3a (n=66)**	**GT1b (n=5)**	**GT2b (n=12)**	**GT4a (n=7)**	**GT4d (n=5)**	**GT5a (n=4)**	**GT6a (n=3)**
24	R	1	-	-	-	-	-	-	-
26	E	-	2	-	-	-	-	-	-
28	T/V/F	1 (T), 4 (V)	2 (V)	-	-	-	-	-	33 (F)
30	H/R/K/T/S/L	3 (H)	3 (K), 2 (T)	100 (R)	100 (K)	86 (L), 14 (R)	100 (R)	-	100 (R)
31	M	-	-	-	50	100	100	-	-
54	H	100	-	40	-	100	100	-	100
93	S/H	1 (S)	3 (H)	-	-	14 (H)	-	-	-

*Only variants detected at frequency >10% of the quasispecies are shown

Table 4.

Prevalence of NS5B resistance associated variants detected in this study*.

		Prevalence of RAVs [%]							
NS5B Position	**Variants Detected**	**GT1a (n=74)**	**GT3a (n=66)**	**GT1b (n=5)**	**GT2b (n=12)**	**GT4a (n=7)**	**GT4d (n=5)**	**GT5a (n=4)**	**GT6a (n=3)**
15	G	-	-	-	100	-	-	-	-
142	T	-	8	-	-	-	-	-	-
159	F	1	-	-	-	-	-	-	-
289	I/L	-	-	-	16.7 (I)	-	-	-	-
316	Y	1	-	-	-	-	-	-	-
392	I	-	-	-	100	-	-	-	-
414	L/V	-	2 (L)	-	-	100 (V)	100 (I)	-	-
419	I	-	98	-	100	100	100	-	100
421	V	15	98	-	92	100	100	50	100
423	I	-	-	-	-	-	-	100	-
426	V	1	-	-	-	-	-	-	-
445	F	-	100	-	100	100	100	100	100
448	C	-	2	-	-	-	-	-	-
482	L	3	100	-	100	100	100	-	100
494	I/A	1 (I)	-	-	100 (A)	-	-	-	-
496	S	-	2	-	-	-	-	-	-
499	A	96	98	-	100	100	100	100	100
553	V	3	100	-	100	100	100	100	-
554	D	1	-	-	-	nd	nd	-	-
556	G	5	100	20	100	nd	nd	-	-

*Only variants detected at frequency >10% of the quasispecies are shown. nd; not determined.

Table 5.

Geographical distribution of HCV RAVs for GT1a and GT3a samples in this study*.

RAV Position	Variants	Genotype 1a			Genotype 3a		
		Aus (n=24)	US/CA (n=48)		Aus (n=33)	US/CA (n=33)	
		Prevalence [%]		P Value	Prevalence [%]		P Value
NS3							
36	L	0	0	NA	100	100	NA
43	S	0	0	NA	3	0	1.000
54	S	0	2	1.000	0	0	NA
55	A	0	4	0.534	0	0	NA
80	K	4	42	<0.001	3	0	1.000
80	R	0	2	1.000	0	0	NA
132	V	0	0	NA	3	0	1.000
155	K	0	2	1.000	0	0	NA
168	H	0	0	NA	0	3	1.000
168	Q	0	0	NA	100	97	1.000
175	L	100	100	NA	100	100	NA
NS5A							
24	R	0	2	1.000	0	0	NA
26	E	0	0	NA	0	3	1.000
28	T	0	2	1.000	0	0	NA
28	V	8	2	0.269	3	0	1.000
30	H	0	4	0.543	0	0	NA
30	K	0	0	NA	3	3	1.000
30	T	0	0	NA	0	3	1.000
54	H	100	96	0.543	0	0	NA
93	S	0	2	1.000	0	0	NA
93	H	0	0	NA	6	0	0.492
NS5B							
142	T	0	0	NA	0	15	0.053
159	F	4	0	0.342	0	0	NA
316	Y	4	0	0.342	0	0	NA
414	L	0	0	NA	0	3	1.000
419	I	0	0	NA	100	97	1.000
421	V	24	10	0.170	100	97	1.000
426	V	4	0	0.342	0	0	NA
445	F	0	0	NA	100	100	NA
448	C	0	0	NA	3	0	1.000
482	L	4	2	1.000	100	100	NA
494	I	0	2	1.000	0	0	NA
496	S	0	0	NA	3	0	1.000
499	A	92	98	0.269	97	100	1.000

RAV Position	Variants	Genotype 1a			Genotype 3a		
		Aus (n=24)	US/CA (n=48)		Aus (n=33)	US/CA (n=33)	
		Prevalence [%]		P Value	Prevalence [%]		P Value
553	V	4	0	0.342	100	100	NA
554	D	0	2	1.000	0	0	NA
556	G	8	2	0.269	100	100	NA

* Only variants detected at frequency >10% of the quasispecies are shown